



Revue internationale d'éducation de Sèvres

40 | décembre 2005
L'éducation dans le monde

Présentation

La généralisation de l'évaluation contribue-t-elle à améliorer la qualité de l'enseignement ?

Pierre-Louis Gauthier



Édition électronique

URL : <http://journals.openedition.org/ries/1228>

DOI : 10.4000/ries.1228

ISSN : 2261-4265

Éditeur

Centre international d'études pédagogiques

Édition imprimée

Date de publication : 1 décembre 2005

Pagination : 25-26

ISSN : 1254-4590

Référence électronique

Pierre-Louis Gauthier, « Présentation », *Revue internationale d'éducation de Sèvres* [En ligne], 40 | décembre 2005, mis en ligne le 17 novembre 2011, consulté le 20 avril 2019. URL : <http://journals.openedition.org/ries/1228> ; DOI : 10.4000/ries.1228

La généralisation de l'évaluation contribue-t-elle à améliorer la qualité de l'enseignement ?

Pierre-Louis Gauthier

Depuis les années soixante-dix, évaluation et éducation forment un couple indissociable à travers les objectifs de l'évaluation traditionnelle des acquis des élèves et des compétences des enseignants. Un couple ancien qui a souvent généré aigreur, voire hostilité.

Depuis, les demandes nouvelles en éducation – demande quantitative avec la massification des filières, demande qualitative avec des études plus longues et plus qualifiantes –, se sont accompagnées d'exigences accrues en matière d'évaluation. Les études de cas présentées ici, portant sur cinq pays (Espagne, Allemagne, Royaume-Uni, France, Sénégal), permettent un examen comparé de ces procédures nouvelles et des problématiques qu'elles engendrent.

L'évaluation en matière d'éducation s'est imposée progressivement dans une conjoncture favorisée par les premières grandes enquêtes internationales qui elles-mêmes emboîtaient le pas aux rapports américains des années soixante-dix. La Catalogne, comme maints autres pays, a depuis inscrit l'évaluation dans sa loi-cadre en matière d'éducation.

Les finalités de l'évaluation englobent des nécessités sociales et des exigences politiques démocratiques de transparence en matière d'éducation. Les considérations économiques sur le coût de l'éducation renforcent la tendance et sont parfois devenues dominantes, comme en France.

Les domaines de l'évaluation varient d'un système à l'autre en fonction des politiques éducatives et de leurs priorités. L'évaluation du domaine pédagogique constitue généralement le premier volet des dispositifs. Les compétences de base des écoliers allemands représentent la première cible des évaluations. À partir de là s'élaborent des réformes fondamentales pour le système éducatif tout entier, comportant notamment de nouveaux standards éducatifs. L'expérience déjà plus ancienne de la Grande-Bretagne concerne l'ensemble du système éducatif, du primaire au secondaire supérieur, mais n'a pas toujours apporté les évolutions positives escomptées. En Catalogne, l'établissement a été choisi comme creuset de l'amélioration de la qualité. En France, l'évaluation devenue systématique et annuelle de deux niveaux « clés » de la scolarité devrait fournir aux établissements des éléments pour un pilotage plus fin de l'action éducative. Le Sénégal a gardé un système d'évaluation qui ne

dépasse guère les résultats obtenus aux examens, ce qui ne favorise pas une évaluation intégrée au processus des apprentissages. On peut ainsi se demander, en négatif en quelque sorte, quels aspects de l'éducation qui ne sont jamais évalués. Ainsi sont mises en relief les lacunes de toute évaluation face à l'acte éducatif tout pétri de culturel et de comportement individuel.

L'examen des méthodes d'évaluation permet d'aborder, au-delà des objectifs, les échelles de l'évaluation et ses modalités. En Allemagne, en France, en Catalogne, ont été créés des organismes nationaux à la fois promoteurs et régulateurs des évaluations nationales. En Grande-Bretagne, la centralisation des tests d'évaluation est toujours fort discutée. De nouveaux processus, telle l'approche systémique ou de nouveaux indicateurs, semblent mieux répondre dorénavant à des contextes changeants permettant d'aborder, après l'évaluation sectorielle, une évaluation plus globale du système (Catalogne, France).

Le moment de l'évaluation ainsi que le choix des évaluateurs doivent se situer avec pertinence dans le temps si particulier de l'éducation. Ainsi, l'évaluation catalane intervient sans ambiguïté en dehors de toute période d'examen. Il en va de même en France.

Enfin, avec la divulgation des résultats, l'accueil des évaluations varie considérablement d'une culture à l'autre. On sait l'émotion ressentie en Allemagne en 2004, lorsque les résultats de PISA placèrent l'école allemande dans le peloton de queue des pays de l'OCDE. Par contre, en France, ces résultats tombèrent dans une certaine indifférence, peut-être parce que le système national d'évaluation et son appareil de diffusion dans le public avaient introduit une certaine accoutumance. On peut déplorer cette indifférence ainsi que l'insuffisante utilisation des évaluations nationales, alors même que le niveau moyen des performances des élèves français est considéré en baisse sensible. Les résultats suscitent toujours la méfiance en Catalogne, cependant que les enseignants britanniques déplorent le surcroît de travail dans les établissements dû à ces opérations d'évaluation. Très variable aussi est l'usage fait des résultats, entre réelle volonté d'amélioration des démarches pédagogiques et discours managérial préoccupé du rendement de l'éducation.

Ces contributions montrent les limites de l'amélioration de l'enseignement découlant uniquement de l'évaluation. Elles convergent généralement vers la nécessité d'associer les partenaires du système éducatif à l'évaluation, grâce à une appropriation des procédures mises en œuvre et à un retour sur les pratiques pédagogiques. En outre, le couple éducation-évaluation, enfin réconcilié, générant une véritable culture de l'évaluation chez les enseignants, se révèle seul capable de s'opposer à la marchandisation de l'éducation et de sauvegarder l'intégrité de l'acte d'enseigner. ■

ÉTUDES DE CAS

Allemagne, Grande-Bretagne,
Catalogne, Sénégal, France

*Allemagne***Une prise de conscience tardive****Hans Konrad Koch**

En matière d'évaluation, l'Allemagne a une tradition beaucoup plus limitée que celle de la France et d'autres pays comparables. Mais ces dernières années, la question de l'évaluation des établissements scolaires a vu, en Allemagne aussi, son importance grandir considérablement, pour deux raisons notamment. On voit maintenant, plus tard que dans d'autres pays, se multiplier les débats sur une plus grande autonomie de l'école. L'accès des établissements scolaires à l'autonomie va de pair avec l'utilisation de l'outil qu'est l'évaluation pour garantir et améliorer la qualité des cours. D'autre part, les résultats des enquêtes internationales sur les résultats scolaires comme TIMSS, PISA et IGLU ont, depuis la fin des années quatre-vingt-dix, toutes remis en question la qualité du système éducatif allemand dans son ensemble. En conséquence, on accorde une grande importance à, notamment, la recherche empirique sur l'enseignement, l'observation permanente sur la base d'indicateurs du système scolaire ainsi que le développement d'une culture de l'évaluation au niveau des établissements.

27

PISA, le déclic pour une nouvelle réforme

Fin 2001, les mauvais résultats obtenus par les élèves allemands à l'enquête PISA ont causé un choc. Experts et hommes politiques savaient depuis longtemps que l'enseignement était une question d'avenir et qu'il fallait réformer le système éducatif. Mais il a fallu attendre ces résultats pour que l'opinion publique prenne conscience des faiblesses du système éducatif et de la nécessité d'une nouvelle réforme de l'enseignement. Cette prise de conscience a été le préalable à une action politique.

On peut résumer ainsi les résultats allemands lors de l'étude PISA 2003 :

- dans les trois domaines de compétences qui ont fait l'objet de l'évaluation, le score global des élèves allemands se trouve au minimum dans la moyenne des trente pays de l'OCDE ; lors de PISA 2000, leurs résultats étaient inférieurs à la moyenne ;
- de façon inchangée, le groupe à risque de ceux qui atteignent au maximum le niveau de compétences le plus bas est de 23 % ;

- l'Allemagne est, parmi tous les pays industriels comparables, le pays avec la plus étroite corrélation entre l'origine sociale et les résultats scolaires;
- comparée aux autres pays de l'OCDE, la dispersion des résultats entre les établissements est très grande.

Objectifs de la réforme

La conséquence de ces mauvais résultats est qu'il existe aujourd'hui, en Allemagne, un consensus sur la nécessité d'une nouvelle et profonde réforme de l'enseignement. L'objectif est d'améliorer durablement la qualité du système éducatif allemand et de réduire enfin l'écart dramatique entre l'origine sociale et les résultats scolaires. La réalisation de ces objectifs suppose un changement d'état d'esprit radical. Il faut passer d'un système éducatif par tradition très sélectif à un système qui encourage les élèves tout en exigeant des résultats, qui accorde la priorité au développement de chaque enfant et de chaque adolescent en s'appuyant sur ses points forts.

Entre-temps, le gouvernement fédéral et les Länder ont pris toute une série de mesures, en particulier dans les domaines suivants :

- la promotion précoce : outre l'augmentation du nombre de crèches selon les besoins, la mission éducative du jardin d'enfants et la qualification des personnels ont la priorité;
- la multiplication des écoles à temps plein : d'ici 2007, le gouvernement fédéral va mettre à la disposition des Länder des crédits d'investissements d'un montant de quatre milliards d'euros pour la multiplication des écoles à temps plein en fonction des besoins; ces nouvelles écoles à temps plein doivent faire émerger une nouvelle culture d'apprentissage et d'enseignement centrée sur le développement individuel de chaque enfant;
- l'amélioration des cours en termes de contenus et de méthodes : des modèles de programmes communs au gouvernement fédéral et aux Länder ont pour objectif de renforcer les compétences en mathématiques et en sciences naturelles, d'améliorer la compréhension de l'écrit, de promouvoir systématiquement les enfants de familles d'immigrants ainsi que de renforcer le développement précoce et individuel;
- l'introduction de *standards* éducatifs nationaux : en 2004-2005, les Länder ont commencé à introduire des *standards* éducatifs nationaux dans les disciplines suivantes : allemand, mathématiques, première langue étrangère et sciences naturelles;
- la professionnalisation des personnels : le succès de la réforme de l'enseignement en dépend; la formation initiale et la formation continue doivent préparer à de nouvelles missions telles que le développement individuel, la coopération avec des partenaires dans le domaine de l'éducation informelle, la mise en place de l'autonomie des établissements scolaires et l'évaluation;

– un rapport national sur le système éducatif : à l'issue de longues années de négociations, le gouvernement fédéral et les Länder se sont mis d'accord sur un rapport commun, régulier et indépendant qui portera sur l'ensemble du système éducatif, y compris l'éducation non formelle et informelle.

Vers la généralisation de l'évaluation

La mise en œuvre et le succès de cette nouvelle réforme de l'enseignement sont inimaginables sans l'évaluation. L'évaluation est un aspect essentiel d'une nouvelle gestion du système qui tient compte, de plus en plus, des résultats (*output*) dans les processus éducatifs.

L'évaluation par des enquêtes internationales comme PISA et IGLU est entre-temps devenue une base incontournable pour le processus de réforme de l'enseignement en Allemagne.

Début 2006, un consortium indépendant de scientifiques va, pour la première fois, présenter un rapport sur l'ensemble du système éducatif allemand qui s'appuiera sur des indicateurs. Ce rapport montrera aux politiques, à l'administration et à l'opinion publique si les mesures prises commencent à avoir de l'effet et dans quels domaines il faut les compléter. À long terme, il est prévu que ce rapport soit soutenu par un dispositif national d'évaluation qui pourrait débiter en 2009 sous la forme d'études séquentielles de cohortes et est censé être compatible avec les enquêtes comparatives internationales.

En 2006 sera graduellement mise en place l'évaluation des nouveaux *standards* éducatifs nationaux. Pour assurer une mise en œuvre de la normalisation et de l'évaluation à un niveau commun à tous les Länder, ces derniers ont créé l'Institut d'assurance de la qualité dans le système éducatif. Parallèlement, quatorze instituts seront créés dans les Länder.

À l'heure actuelle, de grands projets de recherche sont en cours d'élaboration afin d'améliorer les bases de l'évaluation des compétences. L'objectif est de savoir comment définir les compétences et comment les évaluer en s'appuyant sur l'ordinateur et l'Internet.



Pour le moment, le terme «évaluation» suscite encore de profondes craintes, en particulier parmi les enseignants. Beaucoup d'entre eux, mais aussi d'élèves et de parents, ne comprennent pas encore la grande chance que l'évaluation interne et externe représente pour l'émergence d'une nouvelle culture de l'apprentissage et de l'enseignement. Mais toute une série d'exemples montrent que, en Allemagne aussi, le développement d'une culture de l'évaluation est lancé. ■

L'évaluation controversée

Margaret B. Sutherland

On sait depuis un certain temps que les examens ou tests traditionnels ne sont pas la seule méthode possible d'évaluation (OCDE 1993). Cependant, en Grande-Bretagne, c'est cette évaluation qui continue à dominer dans le primaire et dans le secondaire.

Dans le primaire

En Angleterre, on a augmenté le nombre des évaluations. La centralisation, suite à l'*Education Reform Act* (1988), a introduit dès 1991 les SAT (*Standard Assessment Tests*) à l'âge de sept, onze et quatorze ans pour évaluer le progrès des élèves en anglais, en mathématiques et en sciences. À seize ans, les élèves passent le GCSE (certificat général d'enseignement secondaire). Le groupe chargé de préparer l'introduction des SAT avait suggéré quelques méthodes innovantes mais on a éliminé la plupart de ces propositions afin d'en simplifier la gestion. Le gouvernement a annoncé en 1999 un objectif à atteindre en 2002 : 80 % des enfants au niveau 4 à l'âge de onze ans. Mais ce but n'a pas été atteint : seuls 75 % ont réussi.

On aurait pu, quand même, conclure que ces tests avaient amélioré la qualité de l'enseignement puisque, depuis 1999, un plus grand nombre d'élèves ont atteint le niveau prescrit. Mais en même temps, on a dû reconnaître qu'un pourcentage significatif n'avait pas bien acquis les compétences de base.

On constate également d'autres inconvénients. Pendant longtemps, les réactions des enseignants n'ont pas été favorables : on s'est plaint du fardeau imposé aux enfants ; on a regretté que la nécessité d'obtenir de bons résultats dans ces trois matières amène à réduire le temps consacré à d'autres disciplines importantes. En outre, les résultats obtenus par les écoles et publiés dans les journaux – les *league tables* – ont suscité des critiques injustes de la part de lecteurs/parents qui ne savent pas bien interpréter ces résultats.

Une autre preuve de la faillite de ces évaluations se trouve au Pays de Galles où l'on a suivi le même système qu'en Angleterre. Quand la création de l'Assemblée nationale de Galles a encouragé une politique indépendante en éducation, on a décidé progressivement d'éliminer les SAT.

En Écosse, le système de tests nationaux est beaucoup moins rigide. Au primaire, les enseignants peuvent faire passer aux élèves les tests en anglais et en mathématiques au moment où ils les estiment prêts ; on offre même aux enseignants un choix de tests dans ces matières. Ainsi, les enseignants peuvent

savoir si leur évaluation personnelle est fondée. En outre, un programme d'évaluation (AAP) organise des tests nationaux (en anglais, en mathématiques et en sciences) pour un échantillon de classe d'âge tous les trois ans. L'inspection peut ainsi savoir si le niveau atteint est satisfaisant. Cette évaluation peut contribuer à l'amélioration de l'enseignement.

Dans le secondaire

Bien que l'évaluation au niveau des *GCSE* et *A-levels* en Angleterre soit un peu moins centralisée, des problèmes sont apparus récemment. Certaines modifications des *A-levels* ont fait naître le reproche que ces examens devenaient trop faciles. On questionne aussi la fiabilité des notes. Un rapport récent (Tomlinson, 2004) a bien expliqué qu'il faut réformer ces examens pour encourager les jeunes à poursuivre leurs études, assurer leurs compétences de base, valoriser les études professionnelles, développer les dons individuels. Et, d'après ce rapport, il faudrait réduire le fardeau d'évaluation. On a proposé enfin, au lieu des certificats traditionnels, un diplôme global qui sanctionnerait les compétences de base aussi bien que la performance dans les cours généraux et professionnels.

Évidemment, ce rapport estime que le système d'examens dans le secondaire ne sert pas à améliorer l'enseignement. Mais le ministre chargé de l'éducation en Angleterre a indiqué au Parlement, en février 2005, que le gouvernement voulait conserver le système des *GCSE* et des *A-levels*. Il y aura peut-être des diplômes spécialisés pour ceux qui se sont engagés dans la voie professionnelle et un diplôme général pour les compétences de base.

Évaluation des enseignants

Ce n'est pas seulement l'évaluation des élèves qui pourrait amener l'amélioration de l'enseignement. On trouve à présent, en Angleterre et en Écosse, des propositions pour une évaluation nouvelle des enseignants : on veut savoir s'ils méritent des augmentations de salaire et s'ils essaient d'améliorer leur enseignement (*Continuing Professional Development – CPD*). En Angleterre, on a déjà expérimenté ce genre d'évaluation quand il a fallu décider quels enseignants étaient susceptibles d'atteindre un niveau de salaire plus élevé. Mais cette expérience ne semble guère démontrer qu'une telle évaluation, menée par les *Head Teachers*, encourage les enseignants à améliorer leur enseignement. On ne peut pas affirmer que le système de « *performance related pay* » a les résultats espérés (Wragg, 2004). De même, en Écosse, on ne sait pas encore si les deux voies de CPD, le *Training for Headship* (pour ceux qui veulent entrer dans l'administration) ou les études pour obtenir le statut de *Chartered Teacher* (pour ceux qui veulent rester dans la salle de classe) amènent une amélioration.

En Angleterre, l'*OFSTED* (office chargé de l'inspection des établissements) a proposé en 2005 de faire précéder chaque inspection par un questionnaire invitant les élèves de l'établissement à évaluer le travail de leurs enseignants. Proposition intéressante.

Évaluation autonome

L'évaluation la plus efficace ne serait-elle pas faite par les enseignants eux-mêmes et par tous ceux qui se trouvent impliqués dans la vie de l'établissement? On développe par exemple en Écosse depuis plus de dix ans l'évaluation de l'éthos de l'école (Munn, 2003). Cette évaluation résulte de consultations avec les enseignants, les élèves, les parents, le personnel auxiliaire. On ne sait pas encore si cette évaluation améliore la qualité de l'enseignement mais, du point de vue affectif, les résultats semblent bons.

L'évaluation la plus efficace, c'est peut-être l'évaluation autonome : le mérite des autres évaluations serait de stimuler cette évaluation. ■

Bibliographie

MUNN P. (2003): « Ethos and discipline in the Secondary School », *Scottish Education*, T.G.K. Bryce, W.M. Humes (eds.), Edinburgh University Press.

OCDE (1993): *La Réforme des Programmes Scolaires; L'Évaluation en Question*, Paris.

TOMLINSON Report (2004): 14-19. *Curriculum and Qualifications Reform*, www.14-19reform.gov.uk

WRAGG E.C. et al. (2004): *Performance Pay for Teachers*, RoutledgeFalmer, London and New York.

Catalogne

**Expériences récentes
d'évaluation**

Carme Amorós Basté

L'évaluation du système éducatif

L'évaluation du système éducatif dans son ensemble est mise en œuvre en Catalogne par le Conseil supérieur de l'Évaluation, organisme créé en 1993. Le précédent se trouve dans la *Llei d'Ordenació General de l'Educació* (Loi d'organisation générale de l'Éducation) de 1990, loi espagnole qui incorporait toute une partie consacrée à la qualité de l'éducation. Une référence importante mentionnait l'évaluation comme l'un des facteurs favorisant la qualité et l'INCE (Institut national de la qualité et de l'évaluation) était créé en tant qu'organisme responsable des évaluations du système éducatif au niveau national. Dès lors, certaines Communautés autonomes jouissant de compétences en éducation créèrent leurs propres organismes d'évaluation. C'est le cas du *Consell Superior d'Avaluació*, en coordination avec l'INCE, devenu l'INCSE, pour la mise en œuvre des études au niveau national et qui réalise les enquêtes d'évaluation générales en Catalogne.

Depuis la création du Conseil, diverses évaluations diagnostiques globales ont été menées à bien : pour l'enseignement primaire en 1995, 1999 et 2003, pour l'enseignement secondaire en 1996-2000, et quelques évaluations plus ciblées comme celle de l'enseignement de la langue anglaise chez les élèves du primaire et des collèges, ainsi que chez les élèves de lycée. Pendant la période 2000-2002, une évaluation globale participative intitulée Conférence nationale de l'éducation a également été réalisée. Elle avait pour objectif de mener une évaluation critique de l'application en Catalogne de la *Loi générale de l'Éducation*.

Il faut également noter la participation à différentes études au niveau international, comme le *Pre Primary Project*, le *SITES M2* coordonné par l'IEA et PISA. Dans le cas de l'enquête PISA 2000, la participation s'est limitée à faire partie de l'échantillon national mais pour PISA 2003, l'échantillon a été élargi et les données mises à disposition ont été plus significatives pour la Catalogne.

Du point de vue de l'évaluation du système, il convient aussi de signaler l'élaboration périodique d'un système d'indicateurs d'éducation qui prend comme référence le système d'indicateurs de l'OCDE publié annuellement sous le titre *Regards sur l'éducation*. Il y a quatre catégories d'indicateurs : contexte, ressources, processus et résultats¹.

1. Les différentes études et rapports publiés peuvent être consultés sur le site du Conseil : <http://www.gencat.net/educacio/csda/index.htm>

L'évaluation des établissements

Les établissements sont, réellement, les centres clés pour une éducation de qualité. L'évaluation des établissements a pris de l'importance depuis 1997, lors de la publication en Catalogne des normes qui en réglaient le développement, aussi bien dans les établissements du primaire que dans ceux du secondaire.

L'évaluation des établissements a un double point de mire : l'évaluation interne, réalisée par l'établissement scolaire et l'évaluation externe, réalisée par l'inspection d'éducation. Dans un cas comme dans l'autre sont pris en compte :

- l'organisation et la gestion générale de l'établissement (gestion éducative globale, connexion avec l'environnement et gestion des ressources humaines;
- le cursus, la didactique et les résultats de l'apprentissage chez les élèves.

L'évaluation des compétences de bases

L'évaluation des compétences de bases est une façon de coordonner l'évaluation des établissements et l'évaluation du système. Depuis 2001, on a mis en œuvre l'évaluation de la maîtrise des compétences de base chez les élèves de dix ans (4^e année de l'enseignement primaire) et de quatorze ans (2^e année de collège). Ces années ont été fixées deux ans avant la fin des deux étapes éducatives de l'enseignement obligatoire, afin d'éviter que ce type d'évaluation soit assimilé à une validation finale et de permettre à l'établissement de pouvoir introduire des changements avant que les élèves n'achèvent l'étape.

Le tableau suivant illustre la situation des deux années pendant lesquelles l'évaluation est appliquée :

Enseignement obligatoire

Âge	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16
Année	1 ^{re}	2 ^e	3 ^e	4^e	5 ^e	6 ^e	1 ^{re}	2^e	3 ^e	4 ^e
Étape	Éducation primaire						Éducation secondaire obligatoire			

Afin de bien définir les compétences de base, différentes études ont été commandées à un groupe d'experts qui a élaboré une liste des compétences dans différents domaines : linguistique, mathématiques, technique et scientifique, social, utilisation des TIC, artistique et éducation physique. Ces compétences ont également été soumises à consultation des différents secteurs de la communauté éducative, voire du monde économique. Ces études ont été faites avec l'aide de la FREREF (Fondation des régions européennes pour la recherche en éducation et en formation). D'autres équipes de professeurs ont été chargées de l'élaboration des tests.

Le *Consell Superior d'Avaluació* réalise une application externe de ces tests sur un échantillon représentatif des établissements de Catalogne, analyse les résultats

et rédige le rapport. Les résultats de l'échantillon des établissements deviennent un référent de l'ensemble du système et permettent à tous les établissements de Catalogne de comparer leurs propres résultats avec ceux de l'échantillon.

Perspectives d'avenir

Au cours des dernières années, le nombre d'actions évaluatives portant sur les établissements scolaires et sur l'ensemble du système a considérablement augmenté. Les points forts et les points faibles ont été objectivés et surtout, des éléments de comparaison ont été introduits au-delà du nombre d'élèves reçus ou collés lors de chaque année scolaire. Néanmoins, à partir des résultats de l'évaluation, il s'avère difficile de voir comment il faut gérer les processus d'amélioration.

Les établissements perçoivent toujours l'évaluation comme une obligation qui leur donne beaucoup de travail et ne leur permet pas de résoudre les nécessités quotidiennes. Ils expriment également des doutes sur l'usage que l'administration éducative fera de leurs résultats.

Il faudrait sans doute une planification plus globale, cohérente et coordonnée des différentes évaluations : celle du système que réalise le *Consell* et celle des établissements que réalisent l'inspection et les établissements eux-mêmes. Cela permettrait d'économiser les efforts fournis et de prévoir des cycles suffisamment espacés dans le temps, dans les différents types et domaines d'évaluation, de façon à laisser un temps suffisant pour bien connaître les résultats de l'évaluation et introduire les changements nécessaires. C'est dans ce sens qu'on travaille actuellement à un nouveau plan d'évaluation définissant des objectifs et des critères d'intervention communs pour les différentes évaluations suivant un modèle de qualité de l'enseignement.

Le débat reste ouvert sur les savoirs que l'école doit transmettre et les apprentissages que les élèves doivent faire dans un monde globalisé et compétitif. Ces apprentissages concernent les connaissances aussi bien les compétences nécessaires pour apprendre à vivre ensemble. Les évaluations doivent être en cohérence avec les contenus et les objectifs éducatifs et ceci exige de la recherche et une actualisation. Dans ce sens, les enquêtes internationales comme PISA mettent en évidence combien les apprentissages scolaires sont encore loin d'une visée transversale et concrète et comment les conditionnements culturels affectent les réponses des élèves des différents pays.

La qualité du système éducatif exige de pouvoir atteindre une égalité des chances réelle pour tous les élèves sans pour autant renoncer à des niveaux d'excellence. L'évaluation permet de connaître les points forts et les points faibles de la réussite de ce grand objectif et il met probablement davantage en évidence les difficultés qui se rencontrent dans les milieux sociaux les moins favorisés. Il faut donc qu'on puisse également faire connaître les bonnes pratiques des établissements innovants et les politiques éducatives qui peuvent être un support aux établissements qui en ont le plus besoin. ■

Entre tradition et évaluation

Racine Senghor

La première onde de choc sur le système éducatif sénégalais est constituée par les États généraux de l'éducation et de la formation en 1981. Toute la communauté nationale se penche alors sur ce corps atteint pour en comprendre le mal et dégager des pistes et des solutions. Les conclusions issues de ces travaux ont permis quelques corrections. Le système éducatif sénégalais est en pleine mutation depuis le lancement, en avril 2000, du programme EPT (éducation pour tous) par la communauté internationale. Une réflexion avait en effet été engagée deux années auparavant, qui avait abouti à la mise au point d'un Programme décennal de l'éducation et de la formation (PDEF) appuyé par la Banque mondiale et par tous les partenaires financiers de l'École. Il s'agit principalement de promouvoir de manière harmonieuse l'éducation au Sénégal au triple plan de l'accès, de la qualité et de la gestion. C'est donc à une reprise globale du système que les autorités du pays ont procédé du point de vue de la conception puis de la mise en œuvre.

La question de l'évaluation, bien sûr, occupe une place importante en tant que composante fondamentale de la qualité mais aussi en tant qu'instrument de mesure de l'évolution du programme décennal et du système éducatif. En quels termes se pose-t-elle? Comment est-elle abordée? Quels résultats en sont attendus? Quels sont les résultats obtenus?

Du contrôle des acquis des élèves

Il faut d'abord noter que, pour l'élémentaire autant que pour le secondaire, il y a encore des programmes disciplinaires et non un curriculum intégrateur. Celui-ci est en construction pour l'élémentaire mais son élaboration se heurte à diverses difficultés. Fondé sur les valeurs nationales et universelles et sur l'approche par les compétences, il devrait entrer en vigueur prochainement et de façon progressive.

Pour le moment, le contrôle des acquis des élèves se fait, comme traditionnellement, selon la formule des devoirs et autres interrogations écrites ou orales. Des compositions trimestrielles pour l'élémentaire et semestrielles pour le secondaire constituent les temps forts de ce contrôle. Pour le secondaire, elles s'accompagnent des résultats des contrôles continus qui intègrent les résultats finaux.

Des examens et concours sanctionnent la fin des cycles : le certificat de fin d'études élémentaires (CFEE), le concours d'entrée en sixième (organisé en fonction des places disponibles au collège, environ 40 % aujourd'hui), le brevet de fin d'études moyennes (ex BEPC), le baccalauréat.

La situation actuelle révèle des résultats globalement insuffisants, une faible efficacité du système. Les taux de redoublement restent élevés (15 % environ dans l'élémentaire). Les taux de réussite aux examens sont faibles (autour de 40 %). Ces résultats appellent des correctifs afin que le système soit plus performant, plus crédible et plus efficient.

Concernant les redoublements, il a été constaté qu'en dehors des problèmes réels liés aux effectifs pléthoriques, aux absences irrégulières et autres carences, il y a comme « une culture dans laquelle les enseignants considèrent normal de voir redoubler une certaine proportion d'élèves parmi les plus faibles » (CONFEMEN, 1999 ; Eisemon, 1997). Il existe ainsi une grande part d'arbitraire dans les décisions de redoublement.

Les mêmes études révèlent que le nombre d'années-élèves nécessaires pour produire un diplômé est très élevé, du fait justement de ces redoublements. « Parmi les élèves qui entrent en sixième, 83,3 % atteignent la dernière année de ce premier cycle secondaire, 47,2 % entrent dans le deuxième cycle et 34,8 % arrivent jusqu'en terminale. Il faut donc 8,6 années-élèves pour produire un diplômé du premier cycle du secondaire, soit deux fois plus que la durée théorique nécessaire. » Pour un bachelier, il faut 28 années-élèves au lieu des sept ans théoriques.

Il apparaît ainsi que le redoublement est excessivement coûteux, davantage en tout cas que, par exemple, la mise à disposition de manuels scolaires qui, eux, permet d'assurer des succès plus nets. La solution pourrait alors résider à la fois dans la satisfaction des carences diverses, dans la formation des maîtres, dans la reformulation des programmes ou curricula et dans des décisions hardies comme le passage automatique et/ou le redoublement seulement en fin de cycle.

De l'évaluation des apprentissages

Ces conclusions résultent d'enquêtes effectuées à large échelle. Celles-ci s'intéressent de plus en plus aux apprentissages et contribuent à apporter un éclairage particulièrement édifiant dans la conduite du Programme décennal de l'éducation et de la formation. C'est ainsi que plusieurs « structures » travaillent à produire des évaluations. Il s'agit entre autres du PASEC (Programme d'analyse des systèmes éducatifs des pays de la CONFEMEN), du SNERS (Système national d'évaluation des rendements scolaires) et de l'INEADE (Institut national pour le développement de l'éducation).

En 2003, l'INEADE a repris une évaluation des rendements scolaires des élèves au CP et du CE2. Il en ressort que « les objectifs évalués sont loin d'être maîtrisés (...). Environ 60 % des élèves (de CP) atteignent le seuil minimum de maîtrise, toutes disciplines confondues. Moins de 30 % atteignent le seuil désiré (75 %) ».

Pour ce qui est du CE2, « en français, un peu plus de la moitié des élèves atteint le seuil minimum de maîtrise, (contre) les deux tiers en mathématiques. Moins de 12 % atteignent le seuil désiré, toutes disciplines confondues ».

La conclusion est lourde. « Dans toutes les disciplines, et quel que soit le niveau considéré, les résultats sont très faibles et reflètent une qualité des apprentissages très insuffisante. Les élèves, en fin d'étape, sont loin d'atteindre les capacités requises pour entamer des acquisitions nouvelles. »

Auparavant, en 1999, le SNERS et le PASEC avaient travaillé sur « les variables contextuelles », « les mathématiques et les sciences », « le français oral/écrit. » Ces travaux ont été vulgarisés à travers des livrets pédagogiques publiés par l'INEADE.

De même a été menée en 2004 une évaluation des expériences tests du Programme décennal, les « initiatives pour le développement de la lecture à l'école et dans la famille » et « l'introduction des langues nationales à l'école. »

Les évaluations menées jusqu'ici intéressent particulièrement l'élémentaire, mais elles vont s'étendre au secondaire afin d'avoir une vue du système dans sa globalité.

Aujourd'hui, le problème est davantage d'installer, à tous les niveaux du système, une véritable culture de l'évaluation, car une massification inéluctable des effectifs est en cours avec un taux brut de scolarisation voisin de 80 %. Par ailleurs on est, au Sénégal, un peu à la croisée des chemins. Il faut introduire du nouveau sans tout bouleverser de but en blanc ; réformer intelligemment en procédant par touches progressives sur une échelle de dix ans ; garder constamment à l'esprit les principes généraux du PDEF – « promouvoir une école performante où tous les apprenants sont aidés à réussir et à aller le plus loin possible dans les apprentissages » – et accorder une attention particulière à « la mise en place d'un curriculum pertinent, d'un dispositif d'évaluation et de remédiation ». ■

Bibliographie

Annuaire statistique, ministère de l'Éducation, DEPRE, 2004.

Rapport PASEC, CONFEMEN, 1999.

Rapport Eisemon, 1997.

CAILLODS Françoise, *Le financement de l'enseignement secondaire : enjeux et perspectives*, SEIA, Dakar, 2004.

Rapports INEADE, Dakar, 2004.

Rapports SNERS, Dakar, 2004.

Documents du PDEF, ministère de l'Éducation, Dakar

France

Vers une culture de l'évaluation

Alain Michel

C'est à partir des années 1980 que l'évaluation devient une préoccupation majeure aux divers niveaux de responsabilité, depuis le niveau national jusqu'à celui de l'enseignant dans sa classe. Ce phénomène n'est pas spécifique à la France. Il s'inscrit dans un contexte international qui doit largement au rapport « *A nation at risk* » (1986) soulignant l'inefficacité du système éducatif américain.

**Le développement progressif
d'une culture de l'évaluation¹**

La pression de l'opinion publique pour davantage de transparence du fonctionnement de l'école a fait d'autant plus tache d'huile que les États-Unis ont financé de grandes études pour mieux situer le niveau de leurs élèves à l'aune internationale : études du *Education Testing Service* de Princeton et de l'IEA. Ces études ont eu de plus en plus d'écho². Aujourd'hui les résultats des enquêtes menées dans le cadre du programme PISA, sous l'égide de l'OCDE, ont un impact certain, même si celui-ci reste trop faible en France.

Dans le même temps, le développement des dépenses accompagnant le processus de massification dans le secondaire et le supérieur et la nécessité de limiter le déficit public ont conduit à élaborer progressivement de nouveaux outils d'évaluation. En France, au niveau national, sont élaborés à partir de 1990 un ensemble de trente indicateurs (*L'État de l'École*) ainsi qu'un ensemble d'indicateurs régionaux (*Géographie de l'École*). Parallèlement sont élaborés des « indicateurs de pilotage des établissements scolaires » (IPES), la recherche en éducation montrant l'importance de l'effet établissement (*school effect*). En particulier, des indicateurs de résultats s'efforcent d'appréhender la réelle valeur ajoutée des collèges et lycées au delà de résultats bruts aux examens ne prenant pas en compte le niveau des élèves recrutés et leur origine sociale.

À partir de 1989 sont également mis en place des tests nationaux standardisés à trois moments cruciaux du parcours scolaire : CE2 (3^e année de l'enseignement élémentaire), classe de 6^e (début du collège) et classe de seconde (début du lycée). L'originalité de ces tests est d'avoir une finalité diagnostique et formative plus que sommative : ces tests passés en début d'année scolaire ont pour objectif de repérer les principales lacunes des élèves, en termes de connaissances et de méthodes, afin de mieux prendre en compte leur hétérogénéité et remédier aux carences observées par des stratégies pédagogiques adaptées.

1. Michel A., « Vers une nouvelle culture de l'évaluation », *Education et Management*, novembre 1992.

2. Black P., Michel A., *Learning from Pupil Assessment: International Comparison*, UCLA, 1998.

L'évaluation, aide au pilotage³ : développements récents

Si le passage des tests s'est développé puis maintenu en début de CE2 et de 6^e, les enseignants percevant mieux les avantages à en tirer, en revanche, on peut déplorer la quasi disparition des tests de début de seconde, dès lors que ceux-ci n'ont plus été organisés au niveau national mais laissés à l'initiative des lycées.

Par contre, l'utilisation des IPES au niveau des collèges et lycées s'est progressivement développée, à un rythme lent il est vrai. Cette évaluation à partir d'indicateurs s'est accompagnée d'expériences d'audit qualitatif menées par des équipes d'inspecteurs et chefs d'établissement. Cette évaluation externe, conçue comme une aide à l'autoévaluation, est appelée à se développer au fur et à mesure que la prise de conscience de l'importance d'une véritable politique pédagogique au niveau de l'établissement scolaire se répandra parmi les partenaires de la communauté éducative, notamment les enseignants.

Autre axe de progrès depuis quelques années : l'évaluation de l'enseignement au niveau des académies, à raison de trois à cinq académies selon les années, par une équipe d'inspecteurs généraux (administratifs et pédagogiques). À la fin de 2005, la quasi totalité des académies aura été évaluée. De fait, il s'agit davantage d'évaluer les conditions de l'enseignement et les modalités de pilotage aux divers niveaux de responsabilité que la qualité de l'enseignement dans les classes. L'analyse des rapports d'inspection des inspecteurs pédagogiques régionaux ne permet pas en effet de pallier l'insuffisance de l'observation directe des séquences d'enseignement par l'équipe d'évaluation. De manière indirecte, on aborde ainsi le domaine dans lequel les procédures d'évaluation ont le plus de mal à progresser, celui des compétences des enseignants.

L'impact de l'évaluation sur la qualité de l'enseignement

Un diagnostic des atouts et des carences des procédures d'évaluation a été établi en 2001 à la demande du Haut Conseil d'Évaluation de l'École (HCEE) créé en novembre 2000⁴. Il souligne l'insuffisante utilisation des résultats des évaluations et préconise de diversifier et améliorer les procédures d'évaluation à tous les niveaux de responsabilité, tout en développant les comparaisons dans le temps et dans l'espace.

Je souhaiterais mentionner brièvement quelques aspects qui me paraissent cruciaux au vu de ma propre expérience au niveau international et national (quatre ans à la direction de l'évaluation et de la prospective du ministère et onze ans à l'inspection générale).

3. Michel A, « Évaluer pour piloter », *Revue Internationale d'éducation* n° 26, Sèvres, juin 2000.

4. Pair C, *Forces et faiblesses de l'évaluation du système éducatif en France*, rapport au HCEE n° 3, octobre 2001.

L'exploitation des résultats aux tests d'évaluation diagnostique est variable selon les écoles, les collèges et les enseignants, mais elle progresse d'une manière générale, notamment dans l'enseignement élémentaire. Toutefois, l'impact sur le niveau des élèves reste incertain, comme l'attestent la persistance de l'échec scolaire et la difficulté à réduire le nombre d'élèves sortant sans qualification du système éducatif.

En France, on ne tient pas suffisamment compte des résultats aux enquêtes internationales du type PISA. Or, s'il est vrai que ces tests souffrent d'un biais culturel, il est important de constater que les résultats des élèves français sont moins bons qu'il y a quelques années. Cela s'explique en partie du fait que les tests actuels évaluent, davantage que des savoirs académiques, la capacité des élèves à utiliser ceux-ci dans des situations ou des problèmes de la vie courante. Peut-être serait-il judicieux d'au moins se poser la question de savoir si nos programmes d'enseignement ne sont pas trop abstraits. Cela rejoint la préoccupation de Georges Charpak et d'autres lauréats du prix Nobel qui ont conduit à développer des approches plus inductives du type *La main à la pâte*. Une autre information importante est le taux relativement élevé de non réponses des élèves français, qui révèle probablement une crainte de se tromper, elle-même due à une attitude des enseignants laissant trop peu de place à la maïeutique.

D'une manière générale, il y a une insuffisante réflexion des enseignants sur les procédures d'évaluation des acquis des élèves et sur l'hétérogénéité des niveaux d'exigence, des critères et des modalités d'évaluation, y compris au sein d'une même discipline. Sans doute est-ce un facteur de démotivation de nombreux élèves et donc un aspect important de ce que pourrait être une politique pédagogique au niveau d'un établissement scolaire.

Une autre priorité est de mieux prendre en compte l'hétérogénéité, non seulement des élèves, mais aussi des enseignants, ce qui suppose un protocole d'évaluation plus systématique et plus scientifique de leurs compétences pédagogiques et d'en tirer les conséquences. En particulier, le fait de ne plus disposer de véritables manuels synthétisant les connaissances est un handicap pour les élèves qui ont du mal à comprendre les cours de tel ou tel enseignant.



On a assisté, au cours des dernières années, à un développement de l'évaluation à divers niveaux de responsabilité qui permettent de mieux percevoir le fonctionnement et les résultats de l'École. Mais c'est sans doute au niveau des enseignants et de l'évaluation des connaissances et des compétences des élèves que l'on a le moins progressé, surtout en matière de stratégies éducatives et pédagogiques exploitant ces évaluations. Cela explique peut-être le poids encore très important de l'origine sociale des élèves sur les résultats et les carrières scolaires. ■

